Historic, archived document

Do not assume content reflects current scientific knowledge, policies, or practices.

A99.7
F7640
C4

United States
Department
of Agriculture

Forest Service

**Intermountain
Research Station**

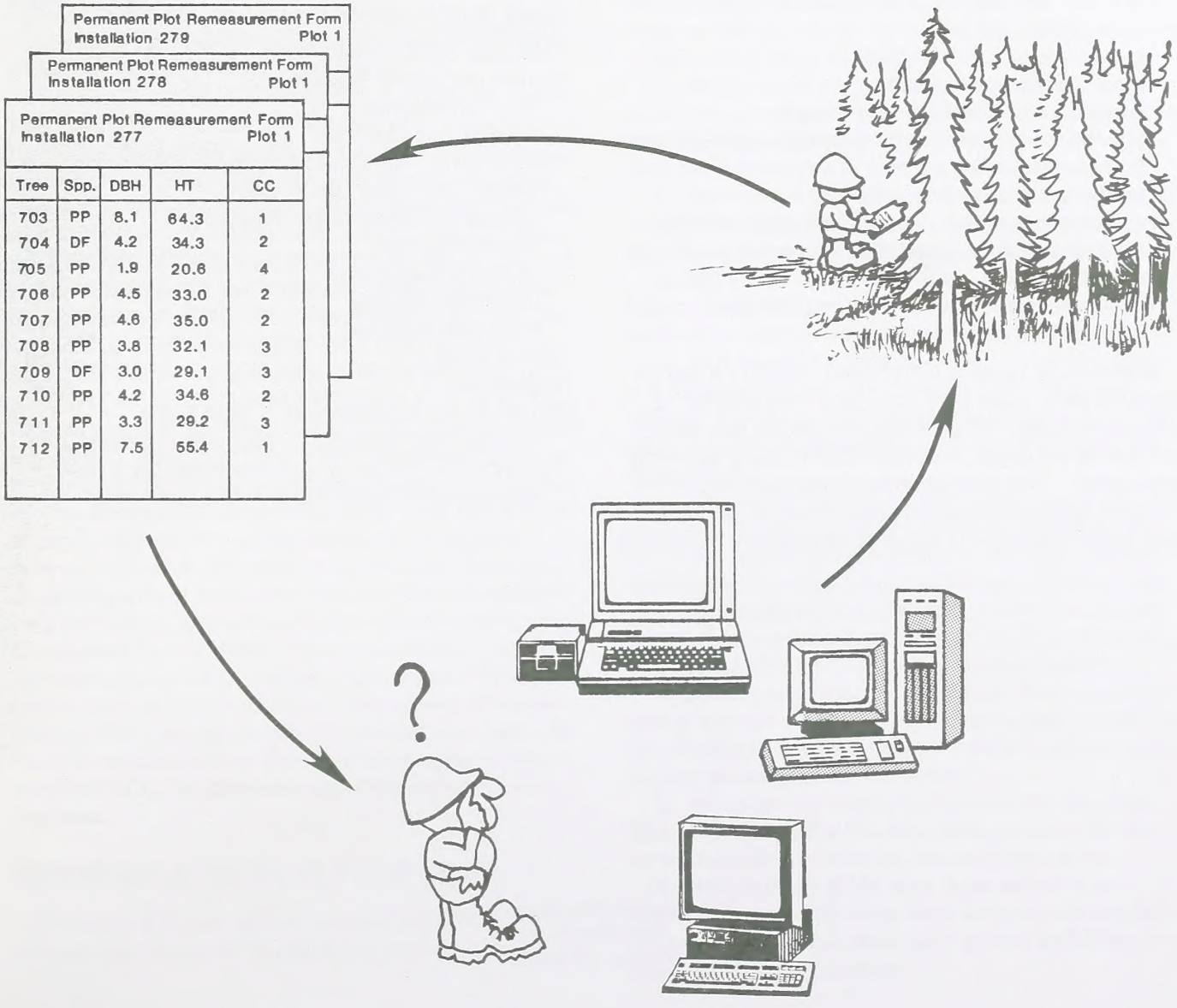
Research Paper
INT-451

February 1992

Managing Data from Remeasured Plots: An Evaluation of Existing Systems

John C. Byrne
Michael D. Sweet

USDA
NAT'L AGRIC. LIBRARY
RECEIVED
MAR 31 '92
CURRENT SERIAL RECORDS
ACQ./ SERIALS BRANCH


Permanent Plot Remeasurement Form
Installation 279 Plot 1
Permanent Plot Remeasurement Form
Installation 278 Plot 1
Permanent Plot Remeasurement Form
Installation 277 Plot 1
Tree Spp. DBH HT CC
703 PP 8.1 64.3 1
704 DF 4.2 34.3 2
705 PP 1.9 20.6 4
706 PP 4.5 33.0 2
707 PP 4.6 35.0 2
708 PP 3.8 32.1 3
709 DF 3.0 29.1 3
710 PP 4.2 34.6 2
711 PP 3.3 29.2 3
712 PP 7.5 55.4 1
?

THE AUTHORS

JOHN C. BYRNE is a research forester with the Intermountain Research Station's Forestry Sciences Laboratory in Moscow, ID.

MICHAEL D. SWEET is a research specialist with the Mission-Oriented Research Program, School of Forestry, University of Montana, Missoula, MT.

RESEARCH SUMMARY

Data from remeasured (or permanent) forest growth plots are a valuable resource to forestry organizations. To fully utilize these data, systems for managing the data are necessary. In 1988, a survey of forestry organizations was conducted to evaluate existing systems for managing remeasured plot data.

For evaluation purposes, 12 desired features of a data base management system (DBMS) for remeasured plot data are described, including:

- import/export large amounts of data
- edit data already in the data base
- error-check data
- store large amounts of data
- store remeasurement data
- safeguard the data and their integrity
- provide for recording commonly measured plot and tree characteristics
- ability to add new data fields
- selective retrieval of data from the data base
- accessibility of the data base software at a reasonable cost
- adequate documentation and help facilities
- analysis/reporting capabilities.

A total of 36 systems were found. Each of these 36 systems were rated as to how close they met the 12 desired features. When features were not fully implemented by a system, descriptions of missing elements were given. This evaluation indicated that most of the systems had some positive characteristics, but none fully implemented all of the desired features. Additional DBMS features, which respondents of the survey noted as desirable, are summarized.

Information from this report can provide guidance to those wishing to improve their management of remeasured plot data.

ACKNOWLEDGMENTS

Major support for this study was provided by the U.S. Department of Agriculture, Forest Service, Timber Management Staff, based in Fort Collins, CO, through a contract with the Inland Northwest Growth and Yield (INGY) Cooperative. We are indebted to the Forest Resources Systems Institute (FORS) for allowing us to use their mailing list to obtain names and addresses of individuals and organizations. We are especially thankful for all those persons who took the time to diligently fill out questionnaires and return them to us. Without their input this study would not have been possible.

CONTENTS

Page

Introduction 1
Desirable System Features 1
Survey of Existing Systems 2
  First Questionnaire 2
  Second Questionnaire 2
Evaluation of Data Base Management Systems 2
  System for Rating DBMS Criteria 2
  Criteria for Rating DBMS Features 2
  System Ratings 5
Results and Discussion 9
  Desirable Features 9
  Additional Features 10
Conclusions 10
References 10
Appendix A—Questionnaire 1: Data Base Management Systems for Remeasured Plots 11
Appendix B—Follow-up Questionnaire: Data Base Management Systems for Remeasured Plots 15

Intermountain Research Station
324 25th Street
Ogden, UT 84401

Managing Data from Remeasured Plots: An Evaluation of Existing Systems

John C. Byrne
Michael D. Sweet

INTRODUCTION

Remeasured (or permanent) forest growth plots are used for a variety of purposes. By a remeasured plot we mean a plot where tagged trees are measured at successive time intervals for monitoring growth and development. Many activities that are integral parts of forest management rely on data from such plots, including the development and validation of forest growth and yield models, the documentation of changes in forest inventory, and the monitoring of forest health. A system for properly managing the data from these plots is essential to their eventual use.

Remeasured plots are typically expensive to install and maintain. Therefore, in recent years, organizations have been formed for sharing existing remeasured plot data or for combining resources to install new plots. One such organization, the Inland Northwest Growth and Yield (INGY) Cooperative, consists of a group of universities, Federal and State forestry agencies, and forest industries in the Northern Rocky Mountains. A major thrust of INGY in recent years has been the consolidation of existing plot data from member organizations into a common data base for use in growth and yield model validation. Utilizing data from these diverse sources has been cumbersome, time consuming, inconsistent, and very costly, because of two major factors: (1) the use of different formats and codes by organizations to store information and (2) the lack of capabilities within organizations to adequately manage the data. We addressed the first factor by designing a data structure that provides a standardized format for the exchange of remeasured plot data (Sweet and Byrne 1990). The second factor, the management of data, is the focus of this report. This paper reports on the results of a 1988 survey of organizations that manage data from remeasured plots. In order to evaluate these data base management systems (DBMS), we defined a set of desirable system features.

DESIRABLE SYSTEM FEATURES

We defined 12 major features that a DBMS for remeasured plot data should have, based on our experiences in managing data from remeasured plots within our own organizations and with cooperators. These features are briefly described below.

1. *Import/export large amounts of data.* The system must allow the input of plot data from existing files (that is, import). For exchange of data, the system must allow the exporting of data to a common format that is acceptable to other computer systems.
2. *Edit data already in the data base.* Data may not be complete or correct when entered into the data base, so the changing of existing data or the addition of more data must be easily done.
3. *Error-check data.* The quality of data is improved if, before it is entered, it is checked for proper coding, reasonableness of values, proper data types, and the uniqueness of variables used as keys.
4. *Store large amounts of data.* Data accumulate rapidly after several measurements of a set of plots, and the system should be capable of handling it.
5. *Store remeasurement data.* All measurements taken on a plot over time should be stored such that successive measurements can be compared.
6. *Safeguard the data and their integrity.* Data are safeguarded by providing for safe, long-term backup, the control of individuals allowed to change data, and the documenting of changes made to the data base.
7. *Provide for recording commonly measured plot and tree characteristics.* Data fields should be provided for recording both common plot-level measurements (for example, sampling design, location, site characteristics, etc.) as well as tree-level measurements (for example, tree number, species, diameter, height, etc.) for more than one measurement.
8. *Ability to add new data fields.* Data requirements change over time, so the data base should have the flexibility to allow new data fields without major reprogramming of the software.
9. *Selective retrieval of data from the data base.* The ability to query the data base permits the user to retrieve data to address a specific question.
10. *Accessibility of the data base software at a reasonable cost.* The data base software should be nonproprietary and reasonably priced and thus enhance availability and use.

11. *Adequate documentation and assistance.* The DBMS, for ease of use, must be well documented and provide assistance for common user problems.

12. *Analysis/reporting capabilities.* Linkage to analysis and reporting software enhances the usefulness of the data. The ability to calculate commonly used stand and plot-level attributes, such as basal area or trees per acre, from the tree data is another important capability. Reports that show incremental changes through time are especially useful.

SURVEY OF EXISTING SYSTEMS

First Questionnaire

A survey of existing systems for managing remeasured plot data was conducted using two separate questionnaires. The first questionnaire was sent to as large an audience of forestry people as possible with the purpose of identifying existing systems and their features.

The first questionnaire (appendix A) was designed to take only a short time to complete. It consisted mainly of questions needing only a yes/no response along with a few short written answers. A cover letter describing the purpose of the survey was sent with each questionnaire. A stamped, self-addressed envelope was enclosed to encourage a response.

The first questionnaire asked if a DBMS was used for remeasured plot data. If the answer was no, the questionnaire was to be returned without proceeding. Otherwise, several questions were asked about the computer hardware and software in use. The next few questions inquired about the amount of data (in terms of plots and trees) presently stored in the system. The next section asked if the system could perform the major tasks previously described. And finally, the respondent was asked to write down other capabilities a DBMS for remeasured plots should include.

Several sources were used in developing the mailing list for the first questionnaire. First, all member organizations of the INGY Cooperative were included. Second, all Regions of the Forest Service were included, as were working units of Forest Service Research Stations that dealt with growth and yield or silviculture. And finally, members from the Forest Resources Systems Institute (FORS), an international nonprofit association that supports and promotes the use of computers in forestry, were contacted. In addition, other members of the forestry profession known to have at least some interest in this field were included on the mailing list. In total, more than 270 organizations were contacted from both the United States and Canada. Questionnaires were sent out in December of 1987, with requests for return by early February 1988. A total of 133 responses were received from this initial mailing.

Second Questionnaire

A second questionnaire (appendix B) was developed for those respondents of the first questionnaire who indicated the existence of a system for managing permanent plot data. This second questionnaire was far more detailed and addressed more specifically the desirable system features that were only touched upon in the first questionnaire. Of the 47 individuals who were mailed this second questionnaire, 36 returned completed questionnaires. The responses received from the second questionnaire were then summarized in tabular form. A rating procedure was developed to evaluate each system's capabilities in regard to the 12 desired DBMS features.

EVALUATION OF DATA BASE MANAGEMENT SYSTEMS

Criteria were developed for evaluating the strengths and weaknesses of each of the 36 systems (those with responses from both the first and second questionnaires). The evaluation was based strictly on answers provided by the organizations responding to the questionnaires. No specific tests have been done to validate actual system components and their adequacy. Nevertheless, we believe that this evaluation provides the necessary information to determine any general trends among the systems as to attributes that are adequately (or inadequately) addressed. Our intention in doing this evaluation was not to pass judgment, either positive or negative, on any of the systems for we realize that organizations have limits on the resources they can expend to implement all of these desired features.

System for Rating DBMS Criteria

For each of the desirable DBMS features described previously, a simple rating of +, 0, 0+, 0–, or – was determined for each system, based on how many of the criteria for that feature were met. Each of the DBMS features had one to three criteria. If there was one criterion, a rating of + was given if the criterion was met or – if the criterion was not met. If there were two criteria, a rating of + was given if both criteria were met, 0 if only one of the criteria were met, or – if neither of the criteria were met. For those features with three criteria, a rating of + was given if all three criteria were met, 0+ if two of the three criteria were met, 0– if one of the three criteria were met, or – if none of criteria were met. Table 1 presents a tabular summary of the rating system.

Criteria for Rating DBMS Features

The criteria used in determining the rating for each DBMS feature and the questions used from

Table 1—Summary of rating system by number of criteria

Number of criteria	Number of criteria met			
	All	Only 2	Only 1	None
	Rating			
1	+			–
2	+		0	–
3	+	0+	0–	–

the questionnaires for evaluating the criteria are described in the following section. The coding system used to reference the questionnaires is as follows: Q1 or Q2 for questionnaire 1 or 2; Roman numeral I, II, or III for the major part of questionnaire 2 where the question is found (not applicable to questionnaire 1); the section number, if applicable, such as A or H; the specific question number, such as 3 or 4. For example, Q2-III-H3 designates questionnaire 2, part III, section H, question 3. In developing the criteria, for summary purposes, we used only those questions that addressed the major DBMS features.

1. *Import/export large amounts of data.*
   - Criteria 1. Capable of importing large amounts of data.
     - "Yes" response to Q1-9.
   - Criteria 2. Capable of exporting large amounts of data.
     - "Yes" response to Q2-III-H7.
2. *Edit data already in the data base.*
   - Criteria 1. Capable of editing data already in the data base.
     - "Yes" response to Q1-10 (or any response to Q2-III-C1).
3. *Error-check data.*
   - Criteria 1. Capable of checking for data errors. Meet all of the following requirements:
     - "Yes" response to either Q1-12 or Q1-13 (error-checking procedure available).
     - "Yes" responses to at least three of the four types in Q2-III-D1 (types of errors procedure checks for).
     - "Yes" response to at least one of the two procedures in Q2-III-D3 (definition of error limits).
     - "Yes" response to Q2-III-D4 (use of previous measurements).
     - "Yes" response to Q2-III-D6 (unique tree numbers).
4. *Store large amounts of data.*
   - Criteria 1. Storage of large amounts of data is not limited.
     - No software categories checked for Q2-III-A3.
5. *Store remeasurement data.*
   - Criteria 1. Experience with storage of remeasurement type data.
     - Number of measurements listed in Q1-7 is greater than 1.
   - Criteria 2. Capable of merging remeasurement data with existing data.
     - "Yes" response to Q1-11.
6. *Safeguard the data and their integrity.*
   - Criteria 1. Capable of safe, long-term backup of data.
     - "Yes" response to Q1-15.
   - Criteria 2. Capable of safeguarding the integrity of the data.

     At least two of the three requirements below are met:
     - "Yes" response to Q2-III-I4 (data base safely available to outside programs).
     - (a) Data entry into data base via transaction file (Q2-III-G3), *or* (b) data entered directly into data base (Q2-III-G3) and a + rating for error-checking data (feature 3 above).
     - "Yes" response to Q2-III-G4 (designation of read-only users).
   - Criteria 3. Capable of automatically recording changes made to the data base.
     - Changes either automatically written to a printer or automatically cataloged (Q2-III-F1).
7. *Provide for recording commonly measured plot and tree characteristics.*
   - Criteria 1. Capable of handling the data from many measurements on a plot.
     - Number of measurements listed in Q2-III-E1 is six or greater.
   - Criteria 2. Capable of sufficiently describing sampling designs. Both of the requirements below are met:
     - For Q2-II-1 (sampling levels), plot, subplot, and tree levels checked *and* at least one of the following levels checked: stand, installation, study.
     - "No" response for Q2-II-2 (restricted description of sampling designs) *or* if a "Yes" response

- "No" response for Q2-II-2 (restricted description of sampling designs) *or* if a "Yes" response for Q2-II-2, then there is more than one sampling design that can be described (Q2-II-3).

- Criteria 3. Capable of storing commonly measured plot and tree characteristics. All requirements below are met:
  - "Yes" response to Q2-II-4 (plot location).
  - "Yes" response to Q2-II-5 (site characteristics).
  - "Yes" response to Q2-II-9 (stand disturbances).
  - "Yes" response to Q2-II-10 (measurement dates).
  - "Yes" response to Q2-II-11 (individual tree attributes).
  - "Yes" response to Q1-20 (comments).

8. *Ability to add new data fields.*
   - Criteria 1. Capable of adding new data fields.
     - "Yes" response to Q1-17.
9. *Selective retrieval of data from the data base.*
   - Criteria 1. Capable of retrieving data based on user-defined search criteria.
     - "Yes" response to Q1-18.
   - Criteria 2. Capable of using stand or plot-level data fields as search criteria.
     - "Yes" response to Q2-III-H1.
   - Criteria 3. Capable of using tree-level data fields as search criteria.
     - "Yes" response to Q2-III-H3.
10. *Accessibility of the data base software at a reasonable cost.*
    - Criteria 1. Reasonable cost for the DBMS software.
      - Cost listed in Q2-I-5 is $1,000 or less.
    - Criteria 2. DBMS does not use proprietary software.
      - "No" response to Q1-3.
11. *Adequate documentation and help facilities.*
    - Criteria 1. Level of computer expertise needed to use the DBMS.
      - "Novice" checked for Q1-21.
    - Criteria 2. Available help facilities and support.
      - Both of the following requirements are met: (1) "Yes" response to Q1-22 (user's manual), and (2) a "Yes" response to at least one of the following questions: Q2-III-J2 (help facilities), Q2-III-J3 (user-support services), Q2-III-J4 (sample data set for testing), Q2-III-J5 (tutorial).
    - Criteria 3. File structures are documented.
      - "Yes" response to Q2-III-A2.
12. *Analysis/reporting capabilities.*
    - Criteria 1. Capable of calculating stand and plot summary attributes.
      - One of the following two requirements are met: (1) Stand and plot attributes are initially calculated (Q2-II-7) and "Yes" response to Q2-II-8 (automatic update when changes made in data base), *or* (2) stand and plot attributes are calculated upon request (Q2-II-8).
    - Criteria 2. Available linkage with statistical software.
      - "Yes" response to Q2-III-K1 and a statistical software package is listed for Q2-III-K1.
    - Criteria 3. Capable of generating reports, especially for incremental changes over time.
      - Both of the following requirements are met: (1) "Yes" response to Q2-III-K5 (incremental change report), and (2) a "Yes" response to at least one of the following questions: Q2-III-K3 (standardized reports), Q2-III-K4 (user-defined reports), Q2-III-K6 (data base summaries).

System Ratings

The ratings for each system reviewed are given in table 2. All organizations are given a numbered code for tabulation purposes. These numeric codes are referenced to the organization name in table 3. Some organizations requested anonymity in any publications based on the questionnaires. These organizations have been designated "Organization A," etc., to protect their identity. Whenever a 0, 0+, or 0– rating is given in the table, a footnote describes the part of the criteria that the system did not meet (see table 4 for the definitions of these footnotes). An * in the table indicates that the questionnaire response was insufficient to evaluate the system for that particular DBMS feature.

Table 2—Evaluation of data base management systems for remeasured plots by organization number

Organization No.[1]	DBMS feature[2]					
	Import, export data (1)	Edit data (2)	Error check data (3)	Large data storage (4)	Store remeasured data (5)	Safeguard data (6)
1	*	+	+	+	+	0_f^+
2	+	+	+	−	+	+
3	+	+	+	−	−	0_f^+
4	+	+	+	+	+	0_f^+
5	−	+	*	+	+	*
6	0_b	+	+	+	+	0_f^+
7	+	+	+	+	0_d	*
8	+	+	+	−	+	0_f^+
9	+	+	+	+	+	0_f^+
10	0_a	+	−	+	+	0_f^+
11	0_b	+	−	+	0_d	$0_{e,f}^-$
12	−	+	−	+	0_d	+
13	+	+	−	−	+	0_f^+
14	+	+	+	+	+	0_f^+
15	+	+	+	+	0_c	+
16	0_b	+	−	+	0_d	$0_{e,f}^-$
17	0_b	+	+	+	+	0_f^+
18	+	+	+	+	+	$0_{e,f}^-$
19	+	+	+	+	+	0_f^+
20	+	+	+	+	+	+
21	*	+	−	+	0_d	0_f^+
22	+	+	+	+	+	0_f^+
23	+	+	+	−	+	+
24	+	+	−	+	+	0_f^+
25	*	+	−	+	+	$0_{e,f}^-$
26	+	+	−	+	0_d	0_f^+
27	+	+	−	+	+	0_e^+
28	+	+	−	+	+	0_f^+
29	+	+	−	−	0_c	0_f^+
30	0_b	+	+	+	0_d	*
31	0_b	+	−	−	+	*
32	+	+	−	+	0_d	+
33	+	+	−	−	+	0_f^+
34	+	+	−	+	+	+
35	0_b	+	−	+	0_d	0_e^+
36	0_b	+	−	−	*	0_f^+

Note: +, −, and 0 codes are described in the text. See table 4 for definitions of the subscripts associated with the 0 codes.

[1]See table 3 for the name and location associated with each of the organization numbers.

[2]DBMS features are defined by number in the text.

(con.)

Table 2 (Con.)

Organization No. [1]	DBMS feature[2]					
	Record common plot, tree data (7)	Add new data fields (8)	Selectively retrieve data (9)	Access to software (10)	Adequate documentation, help (11)	Analysis, reporting (12)
1	$0^{-}_{g,h}$	–	+	*	$0^{-}_{m,n}$	–
2	+	+	+	+	$0^{-}_{m,n}$	–
3	0^{+}_{h}	+	0^{+}_{j}	*	$0^{-}_{m,n}$	$0^{-}_{p,r}$
4	+	+	+	–	+	+
5	–	–	–	*	–	0^{+}_{q}
6	0^{+}_{h}	+	+	+	+	0^{+}_{q}
7	0^{+}_{h}	–	0^{+}_{j}	*	0^{+}_{m}	$0^{-}_{p,r}$
8	$0^{-}_{g,h}$	+	+	0_{k}	$0^{-}_{m,n}$	+
9	0^{+}_{g}	+	+	–	0^{+}_{n}	0^{+}_{r}
10	$0^{-}_{g,h}$	–	$0^{-}_{i,j}$	+	0^{+}_{n}	0^{+}_{q}
11	–	–	0^{+}_{j}	*	0^{+}_{n}	$0^{-}_{q,r}$
12	$0^{-}_{g,h}$	–	0^{+}_{j}	*	–	$0^{-}_{q,r}$
13	0^{+}_{g}	+	0^{+}_{j}	*	$0^{-}_{m,n}$	$0^{-}_{p,q}$
14	0^{+}_{g}	+	+	0_{k}	0^{+}_{m}	+
15	–	+	+	0_{k}	+	+
16	0^{+}_{h}	–	–	*	0^{+}_{m}	$0^{-}_{q,r}$
17	*	–	–	0_{k}	+	0^{+}_{q}
18	–	–	+	0_{l}	0^{+}_{n}	+
19	$0^{-}_{g,h}$	+	+	*	0^{+}_{o}	0^{+}_{q}
20	0^{+}_{h}	–	$0^{-}_{i,j}$	+	+	$0^{-}_{p,q}$
21	*	+	+	+	*	0^{+}_{q}
22	+	+	+	+	$0^{-}_{m,n}$	0^{+}_{q}
23	0^{+}_{g}	–	+	*	0^{+}_{o}	+
24	–	+	0^{+}_{j}	0_{k}	+	0^{+}_{q}
25	0^{+}_{h}	+	–	0_{k}	0^{+}_{n}	–
26	$0^{-}_{g,h}$	–	+	*	$0^{-}_{m,n}$	+
27	$0^{-}_{g,h}$	+	$0^{-}_{i,j}$	+	$0^{-}_{m,n}$	$0^{-}_{p,r}$
28	$0^{-}_{g,h}$	+	+	*	$0^{-}_{m,n}$	$0^{-}_{q,r}$
29	–	+	0^{+}_{j}	0_{l}	0^{+}_{m}	0^{+}_{q}
30	–	+	+	+	$0^{-}_{m,n}$	0^{+}_{q}
31	$0^{-}_{g,h}$	–	–	*	–	0^{+}_{q}
32	–	–	–	+	0^{+}_{o}	0^{+}_{q}
33	0^{+}_{h}	–	+	0_{k}	$0^{-}_{m,n}$	0^{+}_{p}
34	0^{+}_{h}	+	0^{+}_{j}	+	$0^{-}_{m,n}$	0^{+}_{q}
35	–	+	–	*	0^{+}_{o}	0^{+}_{q}
36	–	+	0^{+}_{j}	–	$0^{-}_{m,n}$	0^{+}_{p}

Note: +, –, and 0 codes are described in the text. See table 4 for definitions of the subscripts associated with the 0 codes.
[1]See table 3 for the name and location associated with each of the organization numbers.
[2]DBMS features are defined by number in the text.

Table 3—Directory of organizations by number, name, and location

Number	Name	Location
1	Southern Forest Experiment Station	Pineville, LA
2	Pacific Northwest Research Station	Olympia, WA
3	Organization A[1]	
4	Organization B	
5	Organization C	
6	Organization D	
7	Mead Paper	Escanaba, MI
8	North Central Forest Experiment Station	St. Paul, MN
9	British Columbia Ministry of Forests and Lands	Victoria, BC
10	Northeastern Forest Experiment Station	Parsons, WV
11	Organization E	
12	Organization F	
13	MacMillan Bloedel Ltd.	Nanaimo, BC
14	Organization G	
15	Organization H	
16	USDA Forest Service, Northern Region	Missoula, MT
17	Washington Department of Natural Resources	Olympia, WA
18	Organization I	
19	Champion Forest Products Ltd.	Hinton, AB
20	Alberta Forest Service	Edmonton, AB
21	Northeastern Forest Experiment Station	Broomall, PA
22	Pacific Northwest Research Station	Portland, OR
23	Organization J	
24	Organization K	
25	North Central Forest Experiment Station	Grand Rapids, MN
26	Organization L	
27	Organization M	
28	Organization N	
29	Great Southern Paper	Cedar Springs, GA
30	Organization O	
31	Organization P	
32	Pacific Forest Consultants, Inc.	Portland, OR
33	Southern Forest Experiment Station	Monticello, AR
34	Northeastern Forest Experiment Station	Warren, PA
35	USDI Bureau of Indian Affairs	Minneapolis, MN
36	Champion International Corporation	Kingsford, MI

[1]So designated to protect anonymity.

Table 4—Key to subscripted codes used in table 2

Subscripted code	Definition
a	Not capable of importing large amounts of data.
b	Not capable of exporting large amounts of data.
c	No experience with remeasurement data.
d	Not capable of merging remeasurement data with existing data.
e	Not capable of adequately safeguarding the integrity of the data.
f	Not capable of automatically recording changes made to the data base.
g	Not capable of sufficiently describing sampling designs.
h	Not capable of fully storing commonly measured plot and tree characteristics.
i	Not capable of using stand or plot-level data fields as search criteria in retrievals.
j	Not capable of using tree-level data fields as search criteria in retrievals.
k	DBMS software cost is greater than $1,000.
l	DBMS is proprietary.
m	Experienced or expert level of computer knowledge needed to use DBMS.
n	Inadequate help facilities and support available.
o	File structures not documented.
p	Not capable of adequately calculating stand and plot summary attributes.
q	Linkage with statistical software not available.
r	Not capable of adequately generating reports, including reports for incremental changes over time.

RESULTS AND DISCUSSION

Desirable Features

From the evaluation in table 2, it is apparent that none of the surveyed systems fully implemented all of the desirable features. This is not surprising because these features make up an ideal DBMS for remeasured plots.

Most of the systems could import and export data (feature 1). There were a few that had trouble exporting data. All of the systems possessed a technique for editing the data already in the data base (feature 2).

Half of the systems did not meet the error-checking criteria (feature 3). The criteria were purposely designed to be stringent due to the importance we feel this feature has to the quality of data. Most of these systems failed to use previous measurements to check the reasonableness of current measurements. This error-checking technique is one of the most useful we have found for detecting errors. Many of the systems also did not check for unique tree numbers in each plot, which becomes especially important for (1) larger plots with many trees and (2) plots with a large number of new ingrowth trees. Only one of the systems depended strictly on human error-checking as opposed to machine-driven procedures. Though human error-checking can be beneficial, we believe well-designed, machine-coded procedures are more efficient and better detectors of most errors.

Storage of large amounts of data (feature 4) was limited only by hardware constraints (a + rating) for most of the systems. The DBMS software should not limit the amount of data that is potentially storable.

Most of the organizations have experience collecting remeasurement data (feature 5). The criterion most commonly failed was the inability to merge remeasurement data with existing data. This is a very necessary procedure because it is vital to assessing changes in plot summary and tree attributes over time, probably one of the main reasons for using remeasured plots.

All of the systems had a procedure for safe backup of the data base (feature 6). And most adequately protected the integrity of their data base by assuring that only approved persons have direct access to the data and that safe procedures are used to enter new data into the data base. Recording changes made to the data base presented a major problem for many of the systems. Only a few had a way of automatically recording changes; most relied on the person making the change to document it. Automating the procedure guarantees that a listing of changes is maintained.

We suspect that most of the systems were designed to meet the data needs for a specific inventory or research study. No place is this more apparent than with feature 7, the recording of commonly measured plot and tree attributes. Most provided the capability to record some data for at least six measurements, but many had difficulty when it came to fully describing sampling designs and common measurements. Many of the systems were adequate for one specific design, namely their present inventory or research study design, but were of limited use for describing other designs. Many could not record information at the subplot level, the hierarchical level commonly used for collecting small tree measurements (see Byrne and Stage [1988] for a detailed discussion of sampling designs). Of the 27 systems that failed the criteria on storage of common plot and tree characteristics, most of them were not able to record comments (for example, a unique feature about a plot or tree that existing data fields do not address). More than one-third could not describe plot location and/or types of stand disturbance. The inability to geographically reference plot location seems to be a major omission.

Nearly half of the systems do not have the capability to add new data fields (feature 8). Because data needs will often change over the time that remeasured plots are monitored, the ability to add new data fields is a real advantage, possibly preventing a major reprogramming effort in the future.

Selective retrieval of data from the data base (feature 9) is a necessary part of the DBMS when analysis of the data is required. Many of the systems have retrieval capabilities, most often with plot-level data fields as search criteria (for example, retrieving data from all plots with a site index greater than 70, etc.). Using tree-level data fields as search criteria was not possible for many. An example of such a search would be looking for all Douglas-fir (*Pseudotsuga menziesii*) trees greater than 5 inches d.b.h. growing on south-facing slopes. Tree-level searches become more important when the remeasured plots are used as a basis for improving an individual tree growth and yield model.

The next feature (10), accessibility of the data base software at a reasonable cost, was difficult to evaluate for many systems because of the lack of information provided in the questionnaire (* in table 1). Cost of the DBMS software was difficult for many to provide, especially because many of the organizations had developed their DBMS software in-house. Some of the organizations were unsure whether their DBMS software was proprietary. For the systems for which cost and proprietary nature were known, the main reason for failing the criteria was expense of DBMS software.

Adequate documentation and help facilities (feature 11) are present on only about one-third of the systems (defined by a user's manual plus one of the following: interactive help, user-support services, sample data set, or tutorial). In addition, more than half of the systems require a person that is experienced or expert in computer knowledge to use the system. Because of these two attributes it would probably be difficult for a new user to quickly learn how to use many of these

systems. On the positive side, most systems have well-documented file structures.

Most of the systems are able to generate some summary reports based on the data base (feature 12). A majority could also produce a report on incremental changes over time. But few had a direct linkage to a statistical software package. For those that did not have this direct link, many exported the desired data to an outside file in a readable format, and then had their statistical software program utilize this generated data file. Though a direct link would be desirable, this alternative method at least provides a convenient way for the data to be used by external statistical software.

Additional Features

The last question on the first questionnaire asked the person to list any additional capabilities their system needed to improve performance for their applications. Half of those surveyed with existing systems indicated that some additional features would be useful. Of the additional features that were needed, many of them were related to the DBMS features previously described. Most of the needs listed fell into the general categories of ease of use and analysis. Other responses were related to error-checking, storage, and retrieval of data.

The large number of responses related to ease of use suggests that some of the systems are highly dependent on experienced users, potentially limiting the use of these valuable data by those with lesser skills. The number of responses seems to suggest a desire to develop systems that are accessible to a range of users, from novice to highly advanced. The responses related to ease of use include the addition of user-friendly menus for operation, full-screen and interactive editing, easier methods for inputting field data, better data output capabilities, direct access or links to statistical, graphics, and GIS software, the system programmed in the up-to-date version of the base language instead of out-dated earlier versions, and, finally, better PC-based systems instead of systems dependent on mini- or mainframe computers.

A need for an improved ability to analyze the data was emphasized by some of the respondents. Within the DBMS, analysis needs included listing and plotting growth computations on the computer screen and programs for comparing and analyzing data from several measurements together. There was also a need by several users to interface the data base with analysis programs, written in FORTRAN and other languages, that resided outside the DBMS.

Several respondents felt that they needed better error-checking capabilities or error-checking that could be used in an interactive mode. Two of the responses indicated a desire for better storage of data, namely, data stored without repetition or storage of more than one measurement period's data in the same files. In regard to retrieval of data, respondents wished for improved search techniques as well as faster processing of search requests.

CONCLUSIONS

To take full advantage of the investments made in remeasured plots, the data must be carefully managed. The evaluation of existing systems, as well as the description of important features of a DBMS for remeasured plots, provide guidance for those charged with the management of their valuable resource of remeasured plot data. This information can be used to help prioritize improvements needed in existing systems, and also as a framework for the development of new systems.

REFERENCES

Byrne, J. C.; Stage, A. R. 1988. A data structure for describing sampling designs to aid in compilation of stand attributes. Gen. Tech. Rep. INT-247. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 20 p.

Sweet, M. D.; Byrne, J. C. 1990. A standardized data structure for describing and exchanging data from remeasured growth and yield plots. Gen. Tech. Rep. INT-271. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 43 p.

APPENDIX A—QUESTIONNAIRE 1: DATA BASE MANAGEMENT SYSTEMS FOR REMEASURED PLOTS

Your Name: _____

Your Organization: _____

Your Address: _____

_____

Your Phone Number: _____

Your Position Title: _____

1. Do you have a data base management system for handling plot and tree data from remeasured plots?

   Yes ____ No ____

If your ANSWER TO (1) IS YES, PLEASE PROCEED with the rest of the questionnaire. If your ANSWER TO (1) IS NO, PLEASE RETURN TIIIS FORM in the enclosed, self-addressed envelope.

2. Can your data management system handle more than one measurement period (i.e. more than 2 measurements from a plot)?

   Yes ____ No ____

3. Is your system proprietary within your organization and therefore not releasable to other organizations?

   Yes ____ No ____ Not sure ____

4. What is the make and model of the computer that your system uses (i.e., IBM-PC, DEC Microvax)?

   _____

5. What operating system does your system use (i.e., MS-DOS, VAX/VMS)?

   _____

6. Approximately how many permanent plots and associated trees do you now have entered into your system?

   Number of Plots _____

   Number of Trees _____

7. How many times have the permanent plots you have entered into your system been measured?

   Number of Measurements _____

8. Please list the commercial software packages and/or the programming languages used by your system.

   _____

   _____

   _____

   _____

9. Is it possible to import plot and tree data into your system (i.e. data built with an external editor, data from other software packages, data from portable data recorders)?

   Yes ____ No ____

10. Does your system have an editing procedure for use in changing data already in the system?

    Yes ____ No ____

11. Does your system have a procedure for merging/overlaying remeasurement data with data already in the data base?

    Yes ____ No ____

12. Does your system have procedures for error-checking data that is imported into the system?

    Yes ____ No ____

13. Does your system have procedures for error-checking data entered interactively into the system?

    Yes ____ No ____

14. What are the storage limits for data in your system?

    _____

15. Are procedures available for exporting all data in the system to a storage medium for backup?

Yes ____ No ____

16. Does your system have a procedure for maintaining an historical record of changes made to the data base?

Yes ____ No ____

17. Is it possible to add new data fields to the system with little or no reprogramming of the system?

Yes ____ No ____

18. Is it possible to retrieve selected data from the data base according to user-defined search criteria?

Yes ____ No ____

19. Does your system totally reside on one computer system?

Yes ____ No ____

20. Does your system provide the opportunity to include comments about pertinent information not documented in specific data elements?

Yes ____ No ____

21. Indicate the minimum computer expertise (i.e. knowledge of operating systems, programming languages) necessary for the average user to access and utilize your system?

____ Novice (i.e. has a basic understanding of the software and operating system)

____ Experienced (i.e. has a broad understanding of the software and operating system.)

____ Expert (i.e. has in indepth understanding of the software and operating system. Some programming or database skills may be required.)

Comments : _____

_____

22. Does your system have a user's manual which documents all the information needed to use the system?

Yes ____ No ____

23. What additional capabilities are needed in your system to improve its performance for your applications?

APPENDIX B—FOLLOW-UP QUESTIONNAIRE: DATA BASE MANAGEMENT SYSTEMS FOR REMEASURED PLOTS

NAME _____

PART I. SYSTEM SPECIFICATIONS

1. Describe the storage hardware required (i.e. disk drives, tape drives, etc.).

   _____

   _____

   _____

2. Describe the required output devices (i.e. printers, etc.). Do not include devices listed in (1) above.

   _____

   _____

3. Describe other hardware devices that are required and not listed above.

   _____

4. List other computers/operating systems that could run your software with few or no programming changes.

   _____

   _____

5. What is the approximate total purchase cost of software used for your system?

   ____ less than $500 ____ $3,000–$5,000

   ____ $500–$1,000 ____ $5,000–$10,000

   ____ $1,000–$2,000 ____ $10,000+

6. Is your system easily adaptable to regions of the country other than the region for which it was designed?

   Yes ____ No ____

PART II. DATA ELEMENTS IN THE SYSTEM.

1. INGY has defined terms to define the hierarchical structure associated with permanent plot designs. The terms listed below identify sampling levels in this structure that range from most encompassing to least encompassing. Check the hierarchical levels that are represented in your system.

   ____ Study (A grouping of installations or stands designed to answer a specific question.)

   ____ Installation/Compartment (A logical grouping of stands, typically limited to a specific geographic locality.)

   ____ Stand (An area of ground with relatively uniform conditions and a unique treatment combination.)

   ____ Plot (A collection of trees, i.e. a fixed area plot or prism point.)

   ____ Sub-plot (A sub-sampling scheme for a plot.)

   ____ Tree

   ____ Sub-sections of trees (Stem profile information.)

   Others: _____

   _____

2. Is your system restricted to any particular sampling scheme or plot configuration?

   Yes ____ No ____

3. If (2) above is YES, briefly describe the sampling schemes or plot configurations that can be described with your system.

   _____

   _____

4. Do data fields exist in your system for description of plot location (i.e. latitude/longitude, UTM coordinates, legal description, etc.)?

   Yes ____ No ____

5. Do data fields exist in your system for description of site characteristics (i.e. elevation, slope, aspect, ecological habitat type, site index, site productivity class, etc.)?

   Yes ____ No ____

6. Do data fields exist in your system for description of calculated stand attributes (i.e. trees/acre, basal area/acre, quadratic mean diameter, etc.)?

   Yes ____ No ____

7. If your answer to (6) is YES, which option below describes the calculation of stand attributes:

   ____ Stand attributes are initially calculated for each measurement and then stored in data fields.

   ____ Stand attributes are not stored in data fields but are calculated upon request.

   ____ Other method (please describe): _____

   _____

8. If your answer to (6) is YES, are the calculated stand or plot attributes automatically updated when changes are made to the tree data?

   Yes ____ No ____

9. Do data fields exist in your system for description of stand disturbances (i.e. thinning, fire, brush control, site preparation, etc.), both before and after plot establishment?

   Yes ____ No ____

10. Do data fields exist in your system for description of measurement dates?

    Yes ____ No ____

11. Do data fields exist in your system for individual tree attributes (tree number, species, condition, diameter, height, crown measurements, etc.)?

    Yes ____ No ____

12. Is it possible to use open-ended data fields in your system? (i.e. one to many values for an attribute can be stored in one field, possibly delineated by some delimiter such as a comma or space.)

    Yes ____ No ____

13. Does your system include on-line documentation of data field definitions?

    Yes ____ No ____

14. Are attributes provided within your system to document the precision at which tree data is collected? For example, the precision at which height measurements are taken may vary from clinometer measurements, to ocular estimates, to estimates from a dbh/height regression equation, etc. Is such information documented in your system?

    Yes ____ No ____

PART III. COMMON TASKS ASSOCIATED WITH THE MAINTENANCE, USE AND ANALYSIS OF REMEASURED PLOT DATA.

A. Storage of Data Elements

1. Can the type of file structure used by the program be accessed or duplicated by using a standard text editor?

    Yes ____ No ____

2. Is information provided or available that documents the file structure(s) being used, and how information is stored?

    Yes ____ No ____

3. What determines your system's storage limitations for data?

    ____ Hardware

    ____ Commercial Software

    ____ In-house Software

    Other: _____

4. Does your system track or catalog where the data is physically stored?

    Yes ____ No ____

5. Does the user have a choice in the type of storage medium used to store the data?

    Yes ____ No ____

6. Please enclose a sample of and documentation for the file formats used in your system, if this is appropriate and convenient.

B. Adding New Data Fields to the System

If little or no reprogramming of the system is needed for the addition of new data fields, answer the following questions:

1. Indicate the levels in which new data fields can be added to further document the characteristics of the "permanent plots".

   ____ Study

   ____ Installation/Compartment

   ____ Stand

   ____ Plot

   ____ Sub-plot

   ____ Tree

   ____ Sub-sections of trees

   Other: _____

2. Can acceptable codes or limits be defined for data fields added to the system?

   Yes ____ No ____

3. Is there any limit to how many new data fields can be added?

   Yes ____ No ____

4. If (4) above is YES, briefly describe the limitations.

   _____

   _____

C. Editing Data Elements

1. Check the method that best describes the general editing procedure used by your system (otherwise describe the method used by your system).

   ____ Interactive Editing

   ____ Export data, edit, and import it back into the system.

   ____ Both methods listed above.

   Other Methods: _____

   _____

2. If your system can handle stand, plot, and tree-level data, is it possible to edit the stand or plot-level data without also accessing the associated tree data?

   Yes ____ No ____

D. Error-checking of Data Put into the System

If error-checking is done on all new data entered into the system answer the following questions:

1. Check if these items are included during error-checking.

   ____ Proper Data Type (integer, character, etc.)

   ____ Proper Code Usage (acceptable code for a coded data field)

   ____ Reasonable Data (range and computational checking)

   ____ Unique Identification Numbers (indication of duplicate records)

2. Are checks made to see if values for certain data fields fall within a proper range of values?

   Yes ____ No ____

3. If (2) above is YES, are the limits user-defined or pre-programmed?

   ____ User-defined.

   ____ Pre-programmed.

4. Are previous measurements on a tree used to check the reasonableness of the current measurement? (For example, can previous growth be calculated from past measurements and then be used to check the accuracy of present growth rates?)

   Yes ____ No ____

5. When remeasurement data is entered, are trees that have current measurements, but coded as dead or cut at a previous measurement, flagged for further checking?

   Yes ____ No ____

6. Are tree numbers verified as unique within each plot?

   Yes ____ No ____

E. Remeasurement Data

1. How many possible measurement dates can be stored in the data base for any stand or plot?

   _____

2. Do routines exist, either within the system or in conjunction with the system, for generating tally sheets for use in remeasurement?

   Yes ____ No ____

3. Have you developed routines within your database system to upload and download remeasurement data to field data recorders?

   Yes ____ No ____

4. If (2) and/or (3) above is YES, are at least the most recent past tree measurements provided on the tally sheet or field data recorder for use by the field crew?

   Yes ____ No ____

F. Historical Record of Changes

1. Check the method that best describes how changes made to the data base are documented (otherwise describe the method used by your system).

____ Any change documentation is left up to the user.

____ Any change is automatically written to a printer for subsequent storage.

____ Any change is automatically cataloged in a separate data file.

Other method: _____

_____

2. If changes are automatically cataloged in a data file, who is allowed access to this file?

____ Any System User.

____ Only Selected Users, such as System Managers.

G. Backup and Security of Data

If your system is capable of backup and storage, then answer the following questions:

1. What kind of storage medium is used for backup?

____ 9-track magnetic tape

____ Tape cartridges

____ Floppy diskettes

Other: _____

2. What procedures are presently used at your facility for safeguarding of backup storage media from damage?

_____

_____

3. How is data entered into the master database?

____ Directly into the master database.

____ Into a transaction file, and then, at the user's discretion, into the master database.

4. Is it possible to designate some users as read-only, that is, able to use the database without changing the master database in anyway?

   Yes ____ No ____

H. Retrieval of Data from the Data Base

1. Can stand-level or plot-level data fields be used as search criteria?

   Yes ____ No ____

2. If (1) above is YES, when stand-level or plot-level criteria are used, does the system give the user the option of retrieving the tree data associated with the stands or plots matching the search criteria?

   Yes ____ No ____

3. Can tree-level data fields be used as search criteria?

   Yes ____ No ____

4. When individual tree data are retrieved, does the system provide the user the option of also retrieving the "plot" data associated with each tree found in the search (i.e., the other tree data in the plot or stand from which the "found" tree came from?

   Yes ____ No ____

5. When individual trees are retrieved, does the system give the user the option of retrieving the stand-level or plot-level data associated with the individual tree of interest (i.e., site characteristics)?

6. When data are retrieved from the data base, does your system count the number of installations, stands, plots and/or tree records found that match the search criteria?

   Yes ____ No ____

7. Can retrieval data be output to an external data file in a pre-defined format?

   Yes ____ No ____

I. Link with Exterior Programming Capability

1. If the system is contained on more than one computer, what hardware/software and communications equipment are needed for this link to another computer?

   _____

   _____

   _____

2. Do essential procedures in your system require programming capability outside the main system? (For example, error-checking might be done with a FORTRAN program outside a data-base management system. So, the answer to this question would be YES.)

3. If computer programs external to your database system are used, what language(s) are they written in?

   _____

4. Is the database (or subsets of it) accessible to outside programs without sacrificing the integrity of the database?

   Yes ____ No ____

5. Would you describe your system as having a loose or close tie to a Geographic Information System (GIS)? A loose tie would be having only UTM coordinates or Latitude/Longitude available for plots whereas a close tie would be a direct link to GIS map sheets, showing location of remeasured plots.

   Loose ____ Close ____

J. Training and Ease-of-Use

1. Please estimate the approximate number of training hours required to teach a person to become reasonably proficient at using your system.

   _____

2. Does your system have interactive help facilities to assist users in operating the system?

   Yes ____ No ____

3. Are user-support services provided?

   Yes ____ No ____

4. Is a sample data set provided for learning and testing the system?

   Yes ____ No ____

5. Is a tutorial for learning the system provided?

   Yes ____ No ____

K. Analysis and Reporting

1. Does your database management system interface directly with a statistical package?

   Yes ____ No ____

   If YES, what is the name of that package?

   _____

2. When missing tree data is encountered by analysis programs in your system, how is it handled? (Check appropriate response or describe another method.)

   ____ Tree is assumed to be alive and values are estimated from previous data.

   ____ Tree is assumed to be dead.

   ____ Missing data causes this record to be excluded from the analysis.

   ____ The program ceases to run.

   ____ The results are unpredictable.

3. Does your system have a standardized set of reports that the user can select from?

   Yes ____ No ____

4. Can reports generated from the system be user-defined?

   Yes ____ No ____

5. Can the system generate reports on incremental changes over time (i.e. growth from one measurement date to another)?

   Yes ____ No ____

6. Can the system generate any global summarizations of the database?

   Yes ____ No ____

7. If answer to (6) above is YES, briefly describe the global summarizations available?

   _____

   _____

L. Miscellaneous

1. How long has your system been in operation?

   _____

2. Is the source code provided for all in-house programs that are part of the system?

   Yes ____ No ____

3. Would you approve of us using your company or organization name in a publication or report based on this questionnaire?

   ____ Yes (actual name) ____ No (alias like Company A)

4. In the future, if the Inland Northwest Growth and Yield Cooperative were interested in using or obtaining your system, would your organization be willing to participate in an actual demonstration utilizing a test data set?

   Yes ____ No ____ Maybe ____

★U.S. GOVERNMENT PRINTING OFFICE: 1992 -673-041/A 1025

Byrne, John C.; Sweet, Michael D. 1992. Managing data from remeasured plots: an evaluation of existing systems. Res. Pap. INT-451. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Research Station. 26 p.

Proper management of the valuable data from remeasured (or permanent) forest growth plots with data base management systems (DBMS) can greatly add to their utility. Twelve desired features for such a system (activities that facilitate the storage, accuracy, and use of the data for analysis) are described and used to evaluate the 36 systems found by a survey conducted in 1988. Most of the systems had some positive characteristics, but none fully implemented all of the desired features. Additional DBMS features, which respondents of the survey noted would be desirable, are also summarized.

KEYWORDS: permanent plots, data base management systems, survey

Printed on recycled paper

The Intermountain Research Station provides scientific knowledge and technology to improve management, protection, and use of the forests and rangelands of the Intermountain West. Research is designed to meet the needs of National Forest managers, Federal and State agencies, industry, academic institutions, public and private organizations, and individuals. Results of research are made available through publications, symposia, workshops, training sessions, and personal contacts.

The Intermountain Research Station territory includes Montana, Idaho, Utah, Nevada, and western Wyoming. Eighty-five percent of the lands in the Station area, about 231 million acres, are classified as forest or rangeland. They include grasslands, deserts, shrublands, alpine areas, and forests. They provide fiber for forest industries, minerals and fossil fuels for energy and industrial development, water for domestic and industrial consumption, forage for livestock and wildlife, and recreation opportunities for millions of visitors.

Several Station units conduct research in additional western States, or have missions that are national or international in scope.

Station laboratories are located in:

Boise, Idaho

Bozeman, Montana (in cooperation with Montana State University)

Logan, Utah (in cooperation with Utah State University)

Missoula, Montana (in cooperation with the University of Montana)

Moscow, Idaho (in cooperation with the University of Idaho)

Ogden, Utah

Provo, Utah (in cooperation with Brigham Young University)

Reno, Nevada (in cooperation with the University of Nevada)

USDA policy prohibits discrimination because of race, color, national origin, sex, age, religion, or handicapping condition. Any person who believes he or she has been discriminated against in any USDA-related activity should immediately contact the Secretary of Agriculture, Washington, DC 20250.